Homework 5

Christophe Hunt

February 28, 2017

Contents

1 Page 228: problem 1

Consider a model for the long-term dining behavior of the students at College USA. It is found that 25% of the students who eat at the college's Grease Dining Hall return to eat there again, whereas those who eat at Sweet Dining Hall have a 93% return rate. These are the only two dining halls available on campus, and assume that all students eat at a one of these halls. Formulate a model to solve for the long-term percentage of students eating at each hall.

Table 1: Present - Next State for Dining

		NEXT STATE	
		Grease Dinning Hall	Sweet Dining Hall
PRESENT STATE	Grease Dining Hall	.25	.75
	Sweet Dining Hall	.7	.93

1.1 Model to solve for long-term percentage

$$Grease_{n+1} = .25\,Grease_n + .7\,Sweet_N$$
$$Sweet_{n+1} = .75\,Grease_n + .93\,Sweet_N$$

2 Page 232: problem 1

Consider a stereo with CD player, FM-AM radio tuner, speakers (dual) and power amplifier (PA) components, as displayed with the reliability. Determine the system's reliability. what assumptions are required in your model?

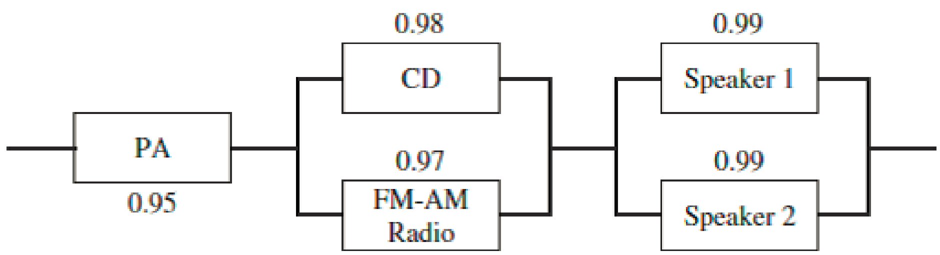


Figure 1: image.

Compenent Reliability

$$R_{s1} = 0.95$$
$$R_{s2} = 0.98 + .97 - (.98 * .97) = 0.9994$$
$$R_{s3} = .99 + .99 - (.99 * .99) = 0.9999$$

Entire system reliability:

$$R_{s1,s2,s3} = .95 * 0.9994 * 0.9999 = 0.9493351$$

3 Page 240: problem 1

Use the basic linear model y = ax + b to fit the following data sets. Provide the model, provide the values of SSE, SSR, SST, and R^2, and provide a residual plot.

```
height <- c(60:80)
weight <- c(132, 136, 141, 145, 150, 155, 160, 165, 170,
            175, 180, 185, 190, 195, 201, 206, 212, 218,
            223, 229, 234)
```

Slope:

```
slope <- function(x,y){
return(((length(x)*sum(x*y)) - sum(x)*sum(y)) /
       ((length(x)*sum(x^2)) - sum(x)^2))
     }
slope(x = height, y = weight)
```

```
## [1] 5.136364
```

Intercept :

```
intercept <- function(x,y){
  (sum(x^2)*sum(y) - sum(x*y)*sum(x)) /
  ((length(x)*sum(x^2)) - sum(x)^2)
}
intercept(x = height, y = weight)
```

```
## [1] -178.4978
```

The linear model y = ax + b for this data set is $y_{weight} = 5.14x_{height} - 178.5$.

Additional measures to aid in our statistical analysis.

Error sum of squares (SSE):

```
SSE <- function(x, y) {
    m <- slope(x = x, y = y)
    b <- intercept(x = x, y = y)
    return(sum((y - (m*x + b))^2))
}
SSE(x = height, y = weight)
```

```
## [1] 24.6342
```

Total Corrected Sum of Squares (SST):

```
SST <- function(x,y) {
    return(sum((y - mean(y))^2))
}
SST(y = weight)
```

```
## [1] 20338.95
```

Regression sum of squares (SSR):

```
SSR <- function(x,y) {
    return(SST(x,y) - SSE(x,y))
}
SSR(x = height, y = weight)
```

```
## [1] 20314.32
```

Coefficient of determination R^2:

```
R2 <- function(x,y){
  return(1 - (SSE(x,y)/SST(x,y)))
}
R2(x = height, y = weight)
```

```
## [1] 0.9987888
```

We can verify the results with the `lm` function in `base R`.

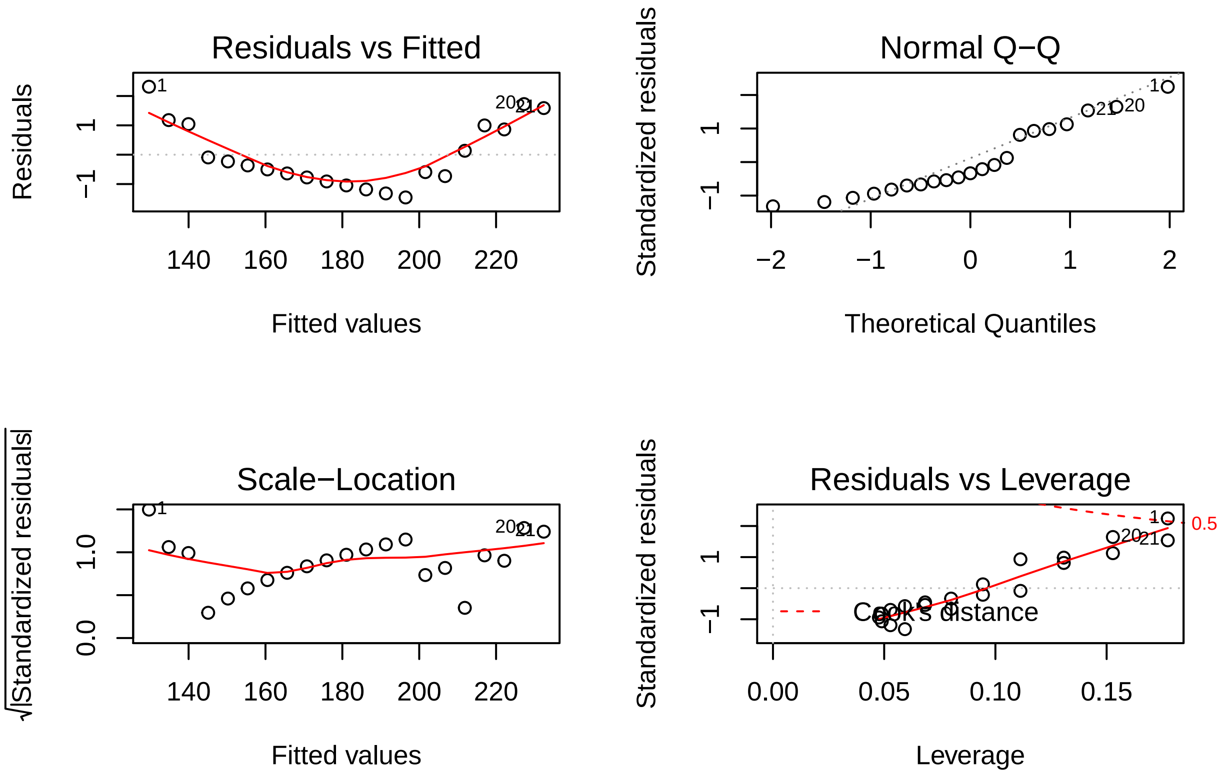
```
library(stargazer)
lm_check <- lm(weight ~ height)
stargazer(lm_check, header = FALSE)
```

Table 2:

	Dependent variable:
	weight
height	5.136^{***}
	(0.041)
Constant	-178.498^{***}
	(2.883)
Observations	21
R^2	0.999
Adjusted R^2	0.999
Residual Std. Error	1.139 (df = 19)
F Statistic	$15{,}668.140^{***}$ (df = 1; 19)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Diagnostic plots

```r
par(mfrow = c(2,2))
plot(lm_check)
```



From the residual plots we can tell that the residuals show constant variance which violates the models assumptions. This model is not actually valid for this data set.

4 Page 240: problem 2

For Table 2.7, predict weight as a function of the cube of the height.

```
height3 <- height^3
```

Slope:

```
slope(x = height3, y = weight)
```

```
## [1] 0.0003467044
```

Intercept:

```
intercept(x = height3, y = weight)
```

```
## [1] 59.4584
```

```
options(scipen=999)
```

The linear model y = ax + b for this data set is $y_{weight} = 0.000347x^3_{height} + 59.46$.

Additional measures to aid in our statistical analysis.

Error sum of squares (SSE):

```
SSE(x = height3, y = weight)
```

```
## [1] 39.86196
```

Total Correct Sum of Squares (SST):

```
SST(y = weight)
```

```
## [1] 20338.95
```

Regression sum of squares (SSR):

```
SSR(x = height3, y = weight)
```

```
## [1] 20299.09
```

Coefficient of determination R^2:

```
R2(x = height3, y = weight)
```

```
## [1] 0.9980401
```

We can verify the results with the lm function in base R.

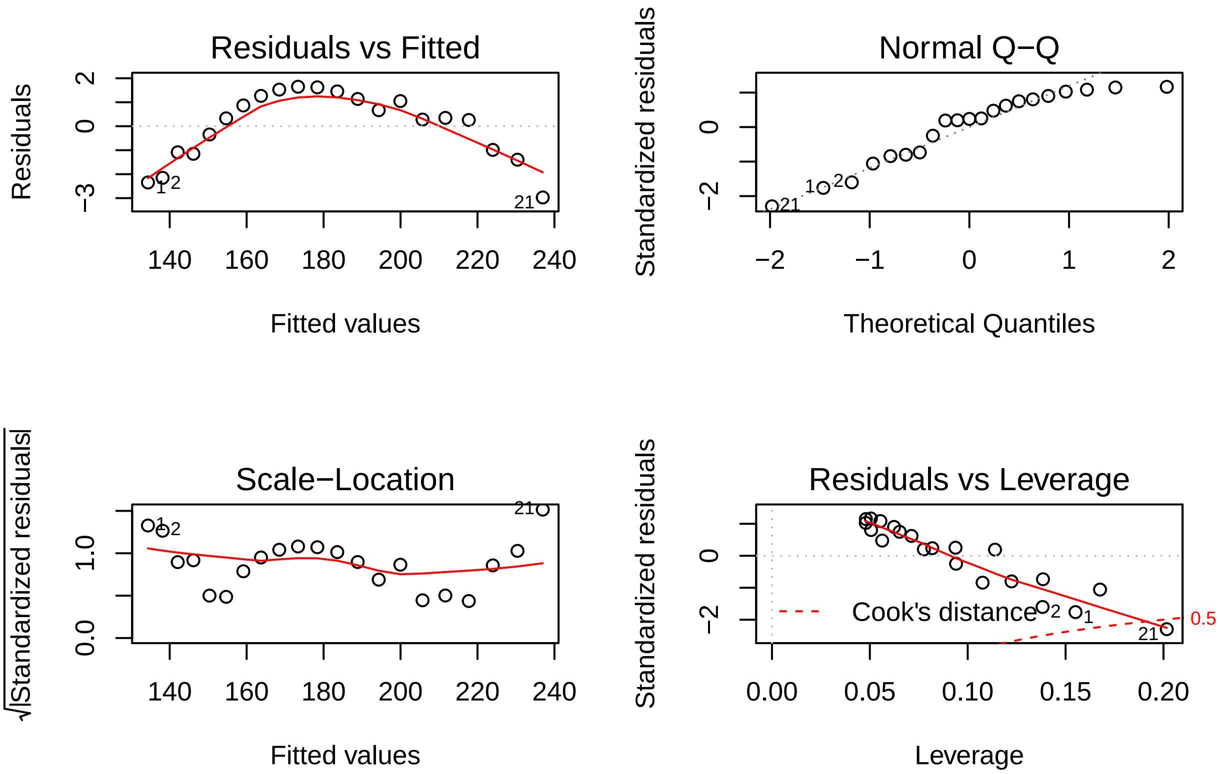
```
library(stargazer)
lm_check <- lm(weight ~ height3)
stargazer(lm_check, header = FALSE)
```

Table 3:

	Dependent variable:
	weight
height3	0.0003***
	(0.00000)
Constant	59.458***
	(1.276)
Observations	21
R^2	0.998
Adjusted R^2	0.998
Residual Std. Error	1.448 (df = 19)
F Statistic	9,675.458*** (df = 1; 19)
Note:	*p<0.1; **p<0.05; ***p<0.01

Diagnostic plots

```
par(mfrow = c(2,2))
plot(lm_check)
```



There appears to be less constant variance in this residual plot, this is the least worst model of the two using a basic linear model.